

RESPONSE

I. Status of the Claims

Claim 6 has been cancelled without prejudice or disclaimer as being drawn to a non-elected invention. Claim 7 has been amended. Claims 2-5 and 7-11 are therefore presently pending in the case. In an attempt to comply with the revised 37 C.F.R. §1.121 and for the convenience of the Examiner the status of the claims is described hereto as **Exhibit A**.

II. Support for the Amended Specification and Claims

Claim 7 has been amended in response to an objection and to further clarify the claim. Amendment of Claim 7 finds support throughout the specification as originally filed with particular support being provided by the original Claim 1 and the Sequence Listing as originally filed.

As the amendments of Claim 7 is fully supported by the specification and claims as originally filed, they do not constitute new matter. Entry therefore is respectfully requested.

III. Rejection of Claims 2-5 and 7-11 Under 35 U.S.C. § 101

The Action first rejects claims 2-5 and 7-11 under 35 U.S.C. § 101 because the claimed invention lacks patentable utility allegedly because the claimed invention is not supported by either a specific and substantial utility or a well established utility.

The Action recognizes that specification describes metalloproteinases (especially zinc metalloproteases of the ADAMTS family) indicates that the sequences of the present invention encode a metalloprotease (in particular ADAMSTS14).

In support of this recognized assertion, Applicants respectfully submit that SEQ ID NO:20 of the present invention is nearly identical to a sequence that is present in the leading scientific repository for biological sequence data (GENBANK). This sequence, GENBANK Accession No. Q8WXS8(information provided as **Exhibit B**) has been annotated by third party scientists *wholly unaffiliated with Applicants* as “ADAMTS-14 precursor (A disintegrin and metalloproteinase with thrombospondin motifs 14) (ADAM-TS 14)”. Thus Applicants’ assertion that the sequences of the

present invention represent a variant of Q8WXS8 is supported by the evidence provided in **Exhibit C**, which contains an amino acid sequence comparison between SEQ ID NO:20 and the amino acid sequence of Q8WXS8. From this comparison, it can be seen that SEQ ID NO:20 shares a greater than 95% homology with the amino acid sequence of Q8WXS8.

Tissues which express the sequences of the present invention were described in the specification (page 3, line 5) as being human spinal cord, lymph node, bone marrow, trachea, mammary gland, skeletal muscle, pericardium, adipose, esophagus, bladder, fetal kidney, and fetal lung cells. The activity of the protein encoded by the sequences of the present invention is described in several publications (Abstracts provided as **Exhibit D**). One is entitled "Characterization of ADAMTS14, a novel member of the ADAMTS metalloproteinase family" by Bolz *et al.*, 2001 (*Biochim Biophys Acta*. 1522:221-5, 2001). A second is entitled "Cloning and characterization of ADAMTS-14, a novel ADAMTS displaying high homology with ADAMTS-2 and ADAMTS-3" by Colige, *et al.*, 2002 (*J Biol Chem*. 277:5756-66, 2002, Epub 2001 Dec 07). The third is entitled "Cloning, expression analysis, and structural characterization of seven novel human ADAMTSs, a family of metalloproteinases with disintegrin and thrombospondin-1 domains" by Cal, *et al.*, 2002 (*Gene* 283:49-62).

Applicants respectfully submit that these publications constitutes evidence that clearly demonstrates that the proteins of the present invention have function and utility that are both accepted by those skilled in the art. As the legal test for utility simply involves an assessment of whether those skilled in the art would find any of the utilities described for the invention to be credible or believable. Clearly those of skill in the art would recognize that molecules that share such amino acid sequence homology would share protein structure and would thus would also share the same function. This constitutes evidence that clearly supports the specifications assertion that SEQ ID NO:19 and 20 encode a known protein (the metalloproteinase ADAMTS14).

Applicants submit that the presently claimed molecules have been shown to encode, as asserted in the specification as filed, ADAMTS-14 precursor (A disintegrin and metalloproteinase with thrombospondin motifs 14) (ADAM-TS 14) which is known to the art. Thus, the present situation directly tracks Example 10 of the Revised Interim Utility Guidelines Training Materials (pages 53-55), which establishes that a rejection under 35 U.S.C. § 101 as allegedly lacking a patentable utility and

under 35 U.S.C. § 112, first paragraph as allegedly unusable by the skilled artisan due to the alleged lack of patentable utility, is not proper when a full length sequence (such as the presently claimed sequence), and has a similarity score greater than 95% to a protein having a known function (such as the 100% identity between the presently claimed sequences and those of the cited protein (Q8WXS8).

Further evidence of utility of the presently claimed polynucleotide, although only one is needed to meet the requirements of 35 U.S.C. § 101 (*Raytheon v. Roper*, 220 USPQ 592 (Fed. Cir. 1983); *In re Gottlieb*, 140 USPQ 665 (CCPA 1964); *In re Malachowski*, 189 USPQ 432 (CCPA 1976); *Hoffman v. Klaus*, 9 USPQ2d 1657 (Bd. Pat. App. & Inter. 1988)), is the specific utility the present nucleotide sequence has in determining the genomic structure of the corresponding human chromosome, for example mapping the protein encoding regions as described in the specification. As evidence supporting Applicants assertions of the specific utility of the sequences of the present invention in localizing the specific region of the human chromosome and identification of functionally active intron/exon splice junctions is the information provided as **Exhibit E**. This is the result of overlaying the sequence of SEQ ID NO:19 of the present invention and the identified human genomic sequence. By doing this, one is able to identify the portions of the genome that encode the present invention. If these regions of the genome are non-contiguous, this is indicative of individual exons. The results of such an analysis indicates that the sequence of the present invention is encoded by at least 24 exons spread non-contiguously along a region of human chromosome 10, which are contained within four BAC clones AL335344.20, AC007484.2, AC069538.10 and AC007484.2. Thus clearly one would not simply be able to identify the 24 or more protein encoding exons that make up the sequence of the present intention from within the large genomic sequence. Nor, would one be able to map the protein encoding regions identified specifically by the sequences of the present invention without knowing exactly what those specific sequences were. Additionally, it should be noted that the human ADAMSTS14 gene also maps to the same region of human chromosome 10 (at approximately 10q2). This further supports Applicant's position that the sequences of the present invention encodes a variant of the human ADAMSTS14.

Therefore, clearly, the present polynucleotide provides exquisite specificity in localizing the specific region of the human chromosome containing the gene encoding the given polynucleotide, a utility not shared by virtually any other nucleic acid sequence. In fact, it is this specificity that makes this

particular sequence so useful. Early gene mapping techniques relied on methods such as Giemsa staining to identify regions of chromosomes. However, such techniques produced genetic maps with a resolution of only 5 to 10 megabases, far too low to be of much help in identifying specific genes involved in disease. The skilled artisan readily appreciates the significant benefit afforded by markers that map a specific locus of the human genome, such as the present nucleic acid sequence.

Only a minor percentage of the genome actually encodes exons, which in turn encode amino acid sequences. The presently claimed polynucleotide sequence provides biologically validated empirical data (*e.g.*, showing which sequences are transcribed, spliced, and polyadenylated) that *specifically* defines that portion of the corresponding genomic locus that actually encodes exon sequence. Equally significant is that the claimed polynucleotide sequence defines how the encoded exons are actually spliced together to produce an active transcript (*i.e.*, the described sequences are useful for functionally defining exon splice-junctions). The Applicants respectfully submit that the practical scientific value of expressed, spliced, and polyadenylated mRNA sequences is readily apparent to those skilled in the relevant biological and biochemical arts. The presently claimed polynucleotide sequence defines a biologically validated sequence that provides a unique and specific resource for mapping the genome.

An additional utility includes the use of the presently claimed polynucleotides on DNA chips. Further, the Action seems to be requiring Applicants to identify the biological role of the nucleic acid or function of the protein encoded by the presently claimed polynucleotides before the present sequences can be used in gene chip applications that meet the requirements of § 101. Applicants respectfully point out that knowledge of the exact function or role of the presently claimed sequence is not required to track expression patterns using a DNA chip. Given the widespread utility of such "gene chip" methods using *public domain* gene sequence information, there can be little doubt that the use of the presently described *novel* sequences would have great utility in such DNA chip applications. The claimed sequence provides a specific marker of the human genome (see evidence below), and that such specific markers are targets for discovering drugs that are associated with human disease. Thus, those skilled in the art would instantly recognize that the present nucleotide sequence would be an ideal, novel candidate for assessing gene expression using, for example, DNA chips, as the specification details. Such "DNA chips" clearly have utility, as evidenced by hundreds of issued U.S. Patents, as

exemplified by U.S. Patent Nos. 5,445,934, 5,556,752, 5,744,305, as well as more recently issued U.S. Patent Nos. 5,837,832, 6,156,501 and 6,261,776. Accordingly, the present sequence has a specific utility in such DNA chip applications. Clearly, compositions that enhance the utility of such DNA chips, such as the presently claimed nucleotide sequence, must also be useful.

Additionally, since only a small percentage of the genome (2-4%) actually encodes exons, which in-turn encode amino acid sequences. Thus, not all human genomic DNA sequences are useful in such gene chip applications. Thus, the present claims clearly meet the requirements of 35 U.S.C. § 101. It has been clearly established that a statement of utility in a specification must be accepted absent reasons why one skilled in the art would have reason to doubt the objective truth of such statement. *In re Langer*, 503 F.2d 1380, 1391, 183 USPQ 288, 297 (CCPA, 1974); *In re Marzocchi*, 439 F.2d 220, 224, 169 USPQ 367, 370 (CCPA, 1971).

Evidence of the “real world” substantial utility of the present invention is further provided by the fact that there is an entire industry established based on the use of gene sequences or fragments thereof in a gene chip format. Perhaps the most notable gene chip company is Affymetrix. However, there are many companies which have, at one time or another, concentrated on the use of gene sequences or fragments, in gene chip and non-gene chip formats, for example: Gene Logic, ABI-Perkin-Elmer, HySeq and Incyte. In addition, one such company, Rosetta Inpharmatics, was viewed to have such “real world” value that it was acquired by large pharmaceutical company, Merck & Co., for substantial sums of money (net equity value of the transaction was \$620 million). The “real world” substantial industrial utility of gene sequences or fragments would, therefore, appear to be widespread and well established. Clearly, persons of skill in the art, as well as venture capitalists and investors, readily recognize the utility, both scientific and commercial, of genomic data in general, and specifically human genomic data. Billions of dollars have been invested in the human genome project, resulting in useful genomic data (see, *e.g.*, Venter *et al.*, 2001, *Science* 291:1304). The results have been a stunning success as the utility of human genomic data has been widely recognized as a great gift to humanity (see, *e.g.*, Jasny and Kennedy, 2001, *Science* 291:1153). Clearly, the usefulness of human genomic data, such as the presently claimed nucleic acid molecules, is substantial and credible (worthy of billions of dollars and the creation of numerous companies focused on such information) and well-established (the utility of human genomic information has been clearly understood for many years). The sequences of

the present invention have particularly specific utility in DNA gene chip based analysis as they have been identified to contain several coding region nucleotide polymorphisms (see above), thus increasing their utility in DNA gene chip based analysis.

Finally, the Examiner is requested to consider the issue of due process. Applicants understanding is that issued United States patents retain a legal presumption of validity which in this case indicates that the inventions claimed in the cited patents are *legally presumed* to be in full compliance with the provisions of 35 U.S.C. sections 101, 102, 103, and 112. Applicants respectfully submit that, absent a change in the law as enacted by Congress and signed by the President, it is improper for the Examiner to hold Applicants' invention to a different legal standard of patentability. Given the rapid pace of development in the biotechnology arts, it is difficult for the Applicants to understand how an invention fully disclosed and free of prior art at the time the present application was filed, could somehow retain *less* utility and be *less* enabled than inventions in the cited issued U.S. patents (which were filed during a time when the level of skill in the art was clearly lower). Simply put, Applicants invention is *more* enabled and retains *at least as much* utility as the inventions described in the claims of the U.S. patents of record. Any argument to the contrary is at best arbitrary and at worst capricious. Absent authority provided by an act of Congress or Executive order, arbitrary or capricious conduct by an administrative office the U.S. government has historically proven to conflict with the provisions of the U.S. Constitution. The Patent Office does not have the authority to rewrite U.S. law. However, the Patent Office does have a Constitutional obligation to administer U.S. law in an unbiased and procedurally consistent manner. That is what the Applicants are respectfully requesting the Examiner to consider in the present matter. As the issued U.S. Patents cited above are presumed to meet all of the requirements for patentability, including 35 U.S.C. §§ 101 and 112, first paragraph, Applicants respectfully submit that the presently claimed polynucleotide must also meet the requirements of 35 U.S.C. § 101.

Thus in summary, Applicants submit that the presently claimed molecules have been shown to encode, as asserted in the specification as filed, ADAMSTS14 isoforms, whose biological function is known to the art. Thus, the present situation directly tracks Example 10 of the Revised Interim Utility Guidelines Training Materials (pages 53-55), which establishes that a rejection under 35 U.S.C. § 101 as allegedly lacking a patentable utility and under 35 U.S.C. § 112, first paragraph as allegedly unusable

by the skilled artisan due to the alleged lack of patentable utility, is not proper when a full length sequence (such as the presently claimed sequence), and has a similarity score greater than 95% to a protein having a known function. Furthermore this response has described a series of additional substantial, specific, credible and well-established utilities for the present invention. Therefore, Applicants submit that as the presently claimed sequence molecules have been shown to have a substantial, specific, credible and well-established utility, the rejection of the claims under 35 U.S.C. § 101 has been overcome. Thus, Applicants respectfully request that the rejection be withdrawn.

IV. Rejection of Claims 2-5 and 7-11 Under 35 U.S.C. § 112, First Paragraph

The Action next rejects claims 2-5 and 7-11 under 35 U.S.C. § 112, first paragraph, as allegedly the claimed invention is not supported by either a specific asserted utility or a well established utility for the reasons set forth above, one skilled in the art clearly would not know how to use the claimed invention.

Applicants respectfully submit that claims 2-5 and 7-11 have been shown to have “a specific, substantial, and credible utility”, as detailed in the section above. Therefore, one skilled in the art would clearly know how to use the claimed invention and Applicants therefore request that the rejection of claims. Therefore, Applicants submit that as the presently claimed sequence molecules have been shown to have a substantial, specific, credible and well-established utility, and thus the rejection of the claims under 35 U.S.C. § 112, first paragraph has been avoided. Thus, Applicants respectfully request that the rejection be withdrawn.

V. Rejection of Claim 7 Under 35 U.S.C. § 112, Second Paragraph

The Action next rejects claim 7 under 35 U.S.C. § 112, second paragraph, as being allegedly being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. While Applicants in no way agree, Claim 7 has been amended to replace the term “drawn” with the term ‘selected’.

Therefore, Applicants submit that amended Claim 7 is clearly definite and particularly points out and distinctly claims the subject matter which applicant regards as the invention and thus the rejection of the claim under 35 U.S.C. § 112, first paragraph has been avoided. Applicants, therefore,

respectfully request that the rejection be withdrawn.

VI. Conclusion

The present document is a full and complete response to the Action. In conclusion, Applicants submit that, in light of the foregoing remarks, the present case is in condition for allowance, and such favorable action is respectfully requested. Should Examiner Moore have any questions or comments, or believe that certain amendments of the claims might serve to improve their clarity, a telephone call to the undersigned Applicants' representative is earnestly solicited.

Respectfully submitted,

August 1, 2003

Date



Lance K. Ishimoto
Agent for Applicants

Reg. No. No. 41,866

LEXICON GENETICS INCORPORATED
(281) 863-3333



24231

PATENT TRADEMARK OFFICE

Exhibit A

Status of Claims in

U.S. Patent Application Ser. No. 09/938,330

1.(cancelled)

2.(original) An isolated nucleic acid molecule comprising a nucleotide sequence that:

- (a) encodes the amino acid sequence shown in SEQ ID NO: 20; and
- (b) hybridizes under stringent conditions to the nucleotide sequence of SEQ ID NO: 19 or the complement thereof.

3.(original) An isolated nucleic acid molecule according to Claim 2 wherein said nucleotide sequence is present in cDNA.

4.(original) An isolated nucleic acid molecule encoding the amino acid sequence presented in SEQ ID NO:20.

5.(original) An isolated nucleic acid molecule encoding the amino acid sequence presented in SEQ ID NO:22.

6.(cancelled)

7.(currently amended) An isolated nucleic acid molecule comprising a nucleotide sequence encoding an amino acid sequence ~~drawn selected~~ from the group consisting of SEQ ID NOS: 20 and 22.

8.(previously presented) An expression vector comprising a nucleic acid sequence of Claim 4.

9.(previously presented) A cell comprising the expression vector of Claim 8.

10.(previously presented) An expression vector comprising a nucleic acid sequence of Claim 5.

11.(previously presented) A cell comprising the expression vector of Claim 10.



NCBI

Protein

PubMed Nucleotide Protein Genome Structure PMC Taxonomy OMIM Books

Search for

Limits Preview/Index History Clipboard Details

Display Show: Send to File

□ 1: Q8WXS8. ADAMTS-14 precurs...[gi:29337086]

BLink, Domains, Links

LOCUS Q8WXS8 1223 aa linear PRI 15-SEP-2003
DEFINITION ADAMTS-14 precursor (A disintegrin and metalloproteinase with thrombospondin motifs 14) (ADAM-TS 14) (ADAM-TS14).
ACCESSION Q8WXS8
VERSION Q8WXS8 GI:29337086
DBSOURCE swissprot: locus AT14_HUMAN, accession Q8WXS8; class: standard.
extra accessions:Q8TE55,Q8TEY8,created: Feb 28, 2003.
sequence updated: Feb 28, 2003.
annotation updated: Sep 15, 2003.
xrefs: gi: [17483853](#), gi: [17483854](#), gi: [19171187](#), gi: [19171188](#), gi: [18874445](#), gi: [18874446](#)
xrefs (non-sequence databases): MEROPSM12.024, GenewHGNC:14899, MIM 607506, InterProIPR001762, InterProIPR001818, InterProIPR002870, InterProIPR001590, InterProIPR000884, InterProIPR008085, InterProIPR006025, PfamPF01562, PfamPF01421, PfamPF00090, PRINTSPR01705, SMARTSM00209, PROSITEPS50215, PROSITEPS00546, PROSITEPS00427, PROSITEPS50214, PROSITEPS50092, PROSITEPS00142
KEYWORDS Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen; Collagen degradation; Repeat; Extracellular matrix; Alternative splicing; Alternative promoter usage.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (residues 1 to 1223)
AUTHORS Bolz,H., Ramirez,A., von Brederlow,B. and Kubisch,C.
TITLE Characterization of ADAMTS14, a novel member of the ADAMTS metalloproteinase family
JOURNAL Biochim. Biophys. Acta 1522 (3), 221-225 (2001)
MEDLINE [21638061](#)
REMARK SEQUENCE FROM N.A. (ISOFORM A).
REFERENCE 2 (residues 1 to 1223)
AUTHORS Cal,S., Obaya,A.J., Llamazares,M., Garabaya,C., Quesada,V. and Lopez-Otin,C.
TITLE Cloning, expression analysis, and structural characterization of seven novel human ADAMTSs, a family of metalloproteinases with disintegrin and thrombospondin-1 domains
JOURNAL Gene 283 (1-2), 49-62 (2002)
MEDLINE [21856482](#)
REMARK SEQUENCE FROM N.A. (ISOFORM A).
TISSUE=Fetal lung
REFERENCE 3 (residues 1 to 1223)
AUTHORS Colige,A., Vandenberghe,I., Thiry,M., Lambert,C.A., Van Beeumen,J., Li,S.W., Prockop,D.J., Lapierre,C.M. and Nusgens,B.V.
TITLE Cloning and characterization of ADAMTS-14, a novel ADAMTS displaying high homology with ADAMTS-2 and ADAMTS-3

JOURNAL J. Biol. Chem. 277 (8), 5756-5766 (2002)
MEDLINE 21839041
REMARK SEQUENCE OF 29-1223 FROM N.A. (ISOFORMS B; C AND D), AND ALTERNATIVE PROMOTER USAGE.
COMMENT -----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. The original entry is available from <http://www.expasy.ch/sprot> and <http://www.ebi.ac.uk/sprot>

[FUNCTION] Has a aminoprocollagen type I activity processing activity in the absence of ADAMTS2. Seems to be synthesized as a latent enzyme that requires activation to display aminoprocollagen peptidase activity.
[SUBCELLULAR LOCATION] Secreted. Associated with the extracellular matrix (By similarity).
[ALTERNATIVE PRODUCTS] Event=Alternative promoter; Comment=2 isoforms, A (shown here) and B, are produced by use of alternative promoters; Event=Alternative splicing; Named isoforms=4; Name=A; IsoId=Q8WXS8-1; Sequence=Displayed; Name=B; IsoId=Q8WXS8-2; Sequence=VSP_006958; Name=C; IsoId=Q8WXS8-3; Sequence=VSP_006958; VSP_005501; Note=Produced by alternative splicing of isoform B; Name=D; IsoId=Q8WXS8-4; Sequence=VSP_005501; Note=Produced by alternative splicing of isoform A.
[TISSUE SPECIFICITY] Expressed in retina and at low levels in brain, lung and placenta. High expression in fetal tissues.
[DOMAIN] The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix (By similarity).
[PTM] The precursor is cleaved by a furin endopeptidase (By similarity).
[SIMILARITY] Belongs to peptidase family M12B.
[SIMILARITY] Contains 1 disintegrin-like domain.
[SIMILARITY] Contains 1 PLAC domain.
[SIMILARITY] Contains 4 TSP type-1 domains.
FEATURES Location/Qualifiers
source 1..1223
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene 1..1223
/gene="ADAMTS14"
Protein 1..1223
/gene="ADAMTS14"
/product="ADAMTS-14 precursor"
/EC_number="3.4.24.-"
Region 1..68
/gene="ADAMTS14"
/region_name="Splicing variant"
/note="Missing (in isoform B and isoform C).
/FTId=VSP_006958."
Region 1..22
/gene="ADAMTS14"
/region_name="Signal"
/note="POTENTIAL."
Region 23..252
/gene="ADAMTS14"
/region_name="Propeptide"
/note="BY SIMILARITY."
Site 109

Region /gene="ADAMTS14"
/site_type="glycosylation"
/note="N-LINKED (GLCNAC...) (POTENTIAL)."
253..1223
Region /gene="ADAMTS14"
/region_name="Mature chain"
/note="ADAMTS-14."
253..460
Region /gene="ADAMTS14"
/region_name="Domain"
/note="METALLOPROTEASE."
368
Site /gene="ADAMTS14"
/region_name="Splicing variant"
/note="G -> GMQG (in isoform C and isoform D).
/FTId=VSP_005501."
398
Site /gene="ADAMTS14"
/site_type="metal-binding"
/note="ZINC (CATALYTIC) (BY SIMILARITY)."
399
Site /gene="ADAMTS14"
/site_type="active"
/note="BY SIMILARITY."
402
Site /gene="ADAMTS14"
/site_type="metal-binding"
/note="ZINC (CATALYTIC) (BY SIMILARITY)."
408
Region /gene="ADAMTS14"
/site_type="metal-binding"
/note="ZINC (CATALYTIC) (BY SIMILARITY)."
461..551
Site /gene="ADAMTS14"
/region_name="Domain"
/note="DISINTEGRIN-LIKE."
475
Region /gene="ADAMTS14"
/site_type="glycosylation"
/note="N-LINKED (GLCNAC...) (POTENTIAL)."
552..607
Region /gene="ADAMTS14"
/region_name="Domain"
/note="TSP TYPE-1 1."
608..729
Region /gene="ADAMTS14"
/region_name="Domain"
/note="CYS-RICH."
730..846
Region /gene="ADAMTS14"
/region_name="Domain"
/note="SPACER."
847..907
Region /gene="ADAMTS14"
/region_name="Domain"
/note="TSP TYPE-1 2."
868
Region /gene="ADAMTS14"
/region_name="Conflict"
/note="Q -> R (IN REF. 2)."

Region 875..878
/gene="ADAMTS14"
/region_name="Domain"
/note="POLY-ARG."
Region 884
/gene="ADAMTS14"
/region_name="Conflict"
/note="Q -> H (IN REF. 2)."
Region 901
/gene="ADAMTS14"
/region_name="Conflict"
/note="C -> S (IN REF. 2)."
Region 908..967
/gene="ADAMTS14"
/region_name="Domain"
/note="TSP TYPE-1 3."
Region 923
/gene="ADAMTS14"
/region_name="Conflict"
/note="C -> Y (IN REF. 2)."
Site 941
/gene="ADAMTS14"
/site_type="glycosylation"
/note="N-LINKED (GLCNAC...) (POTENTIAL)."
Region 968..1022
/gene="ADAMTS14"
/region_name="Domain"
/note="TSP TYPE-1 4."
Region 1024
/gene="ADAMTS14"
/region_name="Conflict"
/note="N -> S (IN REF. 2)."
Site 1027
/gene="ADAMTS14"
/site_type="glycosylation"
/note="N-LINKED (GLCNAC...) (POTENTIAL)."
Region 1063..1093
/gene="ADAMTS14"
/region_name="Domain"
/note="PLAC."
Region 1100..1221
/gene="ADAMTS14"
/region_name="Domain"
/note="PRO-RICH."
ORIGIN
1 maplrallys 1lplhcalca aagsrtpelh lsgklsdygv tvpcstdfrg rflshvvsgp
61 aaasagsmvv dtpptlprhs shlrvarspl hpggtlwpgv vgrhslyfnv tvfgkelhhr
121 lrpnrrlvvp gssviewqedf relfrqplrq ecvytggvtg mpgavaaisn cdglaglirt
181 dstdffiepl ergqkekeas grthvvyrrre avqqewaepd gdlhneafgl gdlnplllglv
241 gdqlgdterk rrhakpgsys ievllvvdds vvrfhgkehv qnyvlmlni vdeiyhdesl
301 gvhinalvr limvgyrqls sliergnpsr sleqvcrwah sqqrqdpsha ehhdhvvflt
361 rqdfgpsgya pvtgmchplr scalnhedgf ssafviahet ghvlgmehdg qngcadets
421 lgsvmaplvq aafhrfhwsr csklelsryl psydcllddp fdpawpppe lpginysmde
481 qcrfdfgsgy qtclafrtfe pckqlwcshp dnpvfcktkk gppldgteca pgkwcfkghc
541 iwkspeqtyg qdggwsswtk fgscsrscgg gvrssrsrscn npspayggrp clgpmfeyqv
601 cnseecpgty edfraqqcak rnsyyvhqna khswwpyepd ddaqkcelic qsadtgdvvf
661 mnqvvhgdtr csyrdpysvc argecvpvgc dkevgsmkad dkcvvcgdn shcrtvkgtl
721 gkaskqagal klvqipagar hiqiealeks phriivvknqv tgsfilnpkg keatsrtfta
781 mglewedave dakeslktsg plpeaiala lppteggprs slaykyvihe dllpligsnn
841 vlleemdtyle walkswapcs kacgggiqft kygcrrrrdh hmvrhlc当地 kkrpkpirrr

901 cnqhpcsqpv wvteewgacs rscgk1gvqt rg1qcllpls ngthkvmpak acagdrpear
961 rpclrvpcpa qwrlgawsqc satcgegiqq rqvvcrtnan slghcegdrp dtvgvcs1pa
1021 cggnhqnstv radvwelgtp egqwvpqsep lhpinkisst epctgdrsvf cqmevldryc
1081 sipgyhrlcc vscikkasgp npgpdpqpts lppfstpgsp lpgpqdpada aeppgkptgs
1141 edhqhgratq lpgaldtssp gtqhpfa1pet pipgaswsis pttppglpwg wtqtptpvpe
1201 dkqgpgedlr hpgts1paas pvt

//

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

Jul 17 2003 11:56:53



Compare Genomic Sequences

EXHIBIT "C"

Page 1 of 3

FASTA searches a protein or DNA sequence data bank
version 3.3t05 March 30, 2000

Please cite:

W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```
/tmp.fastaCAARWaOhY: 1252 aa
>Lex 221 SEQ ID NO:20
vs /tmp.fastaDAASWaOhY library
searching /tmp.fastaDAASWaOhY library
```

1223 residues in 1 sequences

FASTA (3.34 January 2000) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 40, opt: 28, gap-pen: -12/ -2, width: 16
Scan time: 0.067

The best scores are: opt
gi|29337086|sp|Q8WXS8|AT14_HUMAN ADAMTS-14 precur (1223) 8686

>>gi|29337086|sp|Q8WXS8|AT14_HUMAN ADAMTS-14 precursor (1223 aa)
initn: 8683 init1: 7506 opt: 8686

Smith-Waterman score: 8686; 98.611% identity in 1224 aa overlap (1-1224:1-1221)

	10	20	30	40	50	60
Lex	MAPLRALLSYLLPLHCALCXAAGSRTPELHLSGKLSDYGVTVPCSTDFRGRFLSHVVSGP	::::::	::::::	::::::	::::::	::::::
gi 293	MAPLRALLSYLLPLHCALCAAAGSRTPELHLSGKLSDYGVTVPCSTDFRGRFLSHVVSGP	10	20	30	40	50
	70	80	90	100	110	120
Lex	AAASAGSMVVDTPPTLPRHSSHLRVARSPLHPGGTLWPGRVGRHSLYFNNTVFGKELHLR	::::::	::::::	::::::	::::::	::::::
gi 293	AAASAGSMVVDTPPTLPRHSSHLRVARSPLHPGGTLWPGRVGRHSLYFNNTVFGKELHLR	70	80	90	100	110
	130	140	150	160	170	180
Lex	LRPNRLLVPGSSVEWQEDFRELFRQPLRQECSVYTGGVTGMPGAAVAISNCDGLAGLIRT	::::::	::::::	::::::	::::::	::::::
gi 293	LRPNRLLVPGSSVEWQEDFRELFRQPLRQECSVYTGGVTGMPGAAVAISNCDGLAGLIRT	130	140	150	160	170
	190	200	210	220	230	240
Lex	DSTDFFFIEPLERGQQEKEASGRTHVVYRREAVQQEWAEPDGLHNEAFGLGDPNLLGLV	::::::	::::::	::::::	::::::	::::::
gi 293	DSTDFFFIEPLERGQQEKEASGRTHVVYRREAVQQEWAEPDGLHNEAFGLGDPNLLGLV	190	200	210	220	230
	250	260	270	280	290	300
Lex	GDQLGDTERKRRHAKPGSYSIEVLLVVDDSVVRFHGKEHVQNYVLTLMNIVDEIYHDESL	::::::	::::::	::::::	::::::	::::::
gi 293	GDQLGDTERKRRHAKPGSYSIEVLLVVDDSVVRFHGKEHVQNYVLTLMNIVDEIYHDESL	250	260	270	280	290
	310	320	330	340	350	360
Lex	GVHINIALVRLIMVGYRQSLSLIERGNPSRSLEQVCRWAHSQQRQDPHAEHHHDHVVFLT	::::::	::::::	::::::	::::::	::::::
gi 293	GVHINIALVRLIMVGYRQSLSLIERGNPSRSLEQVCRWAHSQQRQDPHAEHHHDHVVFLT	310	320	330	340	350
	370	380	390	400	410	420
Lex	RQDFGPSPGYAPVTGMCHPLRSCALNHEDGFSSAFVIAHETGHVLGMEHDGQGNGCAETS	::::::	::::::	::::::	::::::	::::::

gi 293	RQDFGPGSGYAPVTGMCHPLRSCALNHEDGFSSAFVIAHETGHVLGMEHDGQGNGCAETS					
	370	380	390	400	410	420
	430	440	450	460	470	480
Lex	LGSVMAPLVQAAFHFRFWSRCSKLELSRYLPSYDCLLDDPFDPAWPQPPELPGINYSMDE					
gi 293	LGSVMAPLVQAAFHFRFWSRCSKLELSRYLPSYDCLLDDPFDPAWPQPPELPGINYSMDE					
	430	440	450	460	470	480
	490	500	510	520	530	540
Lex	QCRCDFGSGYQTCLAFRTFEPCKQLWCSPHDNXXFCKTKKGPPLDGTECAPGKWCFKGHC					
gi 293	QCRCDFGSGYQTCLAFRTFEPCKQLWCSPHDNPYFCKTKKGPPLDGTECAPGKWCFKGHC					
	490	500	510	520	530	540
	550	560	570	580	590	600
Lex	IWKSPEQTYGQDGWSSWTKFGSCSRSCGGGVRSRSRSCNNPSPAYGGRXCLGPMFEYQV					
gi 293	IWKSPEQTYGQDGWSSWTKFGSCSRSCGGGVRSRSRSCNNPSPAYGGRPCLGPMFEYQV					
	550	560	570	580	590	600
	610	620	630	640	650	660
Lex	CNSECPGTYEDFRAQQCAKRNXYYVHQNAKHXWVPYEPDDDAQKCELICQSADXGDVVF					
gi 293	CNSECPGTYEDFRAQQCAKRNSYYVHQNAHKHSWVPYEPDDDAQKCELICQSADTGDVVF					
	610	620	630	640	650	660
	670	680	690	700	710	720
Lex	MNQVVHDGTRCSYRDPYVCARGECVPVGCDKEVGSMKADDKGVCGGDNSHCRTVKGTL					
gi 293	MNQVVHDGTRCSYRDPYVCARGECVPVGCDKEVGSMKADDKGVCGGDNSHCRTVKGTL					
	670	680	690	700	710	720
	730	740	750	760	770	780
Lex	GKASKQAGALKLVQIPAGARHIQIEALEKSPHXXVVKNQVTGSFILNPKGKEATSRTFTA					
gi 293	GKASKQAGALKLVQIPAGARHIQIEALEKSPHRIVVKNQVTGSFILNPKGKEATSRTFTA					
	730	740	750	760	770	780
	790	800	810	820	830	840
Lex	MGLEWEDAVEDAKESLKTGPLPEAIAILALPPTEGGPRSSLAYKYVIHEDLLPLIGSNN					
gi 293	MGLEWEDAVEDAKESLKTGPLPEAIAILALPPTEGGPRSSLAYKYVIHEDLLPLIGSNN					
	790	800	810	820	830	840
	850	860	870	880	890	900
Lex	VLLEEMDTYEWALKSWAPCSKACGGGIQFTKYGCRRRDHHMVQRHLCDHKKRPKPIRR					
gi 293	VLLEEMDTYEWALKSWAPCSKACGGGIQFTKYGCRRRDHHMVQRHLCDHKKRPKPIRR					
	850	860	870	880	890	900
	910	920	930	940	950	960
Lex	CNQHPCSQPVWVTEEWGACSRSCGKLGVQTRGIQCLLPLSNGTHKVMPAKACAGDRPEAR					
gi 293	CNQHPCSQPVWVTEEWGACSRSCGKLGVQTRGIQCLLPLSNGTHKVMPAKACAGDRPEAR					
	910	920	930	940	950	960
	970	980	990	1000	1010	1020
Lex	RPCLRVPCPAQWRLGAWSQCSATCGEGIQQRQVVCRNTANSLGHCEGDRPDTVQVCXLPA					

gi|293 RPCLRVPCPAQWRLGAWSQCSATCGEGIQQRQVVCRTNANSLGHC
970 980 990 1000 1010 1020
1030 1040 1050 1060 1070 1080
Lex CGGNHQNSTVRADVWELGTPEGQWVPQSXPLHPINKISSMCAAEPCTGDRSVFCQMEVLD
gi|293 CGGNHQNSTVRADVWELGTPEGQWVPQSEPLHPINKISS---TEPCTGDRSVFCQMEVLD
1030 1040 1050 1060 1070
1090 1100 1110 1120 1130 1140
Lex RYCSIPGYHRLCCVSCIKKASGPNGPDPGPTSLPPFSTPGSPLPGPQDPADAAEPPGKP
gi|293 RYCSIPGYHRLCCVSCIKKASGPNGPDPGPTSLPPFSTPGSPLPGPQDPADAAEPPGKP
1080 1090 1100 1110 1120 1130
1150 1160 1170 1180 1190 1200
Lex TGSEDHQHGRATQLPGALDTSSPGTQHPFAPETPIPGASWSISPTTPGGLPWGWTQTPTP
gi|293 TGSEDHQHGRATQLPGALDTSSPGTQHPFAPETPIPGASWSISPTTPGGLPWGWTQTPTP
1140 1150 1160 1170 1180 1190
1210 1220 1230 1240 1250
Lex VPEDKGQPGEDLRHPGTSLPADLPGRPPEPCHPTGTFTLCVLPRDSQLRGHT
gi|293 VPEDKGQPGEDLRHPGTSLPAASPVT
1200 1210 1220

1252 residues in 1 query sequences

1223 residues in 1 library sequences

Scomplib [version 3.3t05 March 30, 2000]

start: Tue Jul 29 17:17:39 2003 done: Tue Jul 29 17:17:40 2003

Scan time: 0.067 Display time: 2.117

Function used was FASTA



EXHIBIT "D"

J Biol Chem. 2002 Feb 22;277(8):5756-66. Epub 2001 Dec 07.

[Related Articles](#), [Links](#)

Full text article at
www.jbc.org

Cloning and characterization of ADAMTS-14, a novel ADAMTS displaying high homology with ADAMTS-2 and ADAMTS-3.

Colige A, Vandenberghe I, Thiry M, Lambert CA, Van Beeumen J, Li SW, Prockop DJ, Lapierre CM, Nusgens BV.

Laboratory of Connective Tissues Biology, Experimental Cancerology Research Center, Tour de Pathologie (B23/3), University of Liege, B-4000 Liege, Belgium.

The processing of amino- and carboxyl-propeptides of fibrillar collagens is required to generate collagen monomers that correctly assemble into fibrils. Mutations in the ADAMTS2 gene, the aminopropeptidase of procollagen I and II, result in the accumulation of non-fully processed type I procollagen, causing human Ehlers-Danlos syndrome type VIIC and animal dermatosparaxis. In this study, we show that the aminopropeptide of type I procollagen can be cleaved *in vivo* in absence of ADAMTS-2 activity and that this processing is performed at the cleavage site for ADAMTS-2. In an attempt to identify the enzyme responsible for this alternative aminoprocollagen peptidase activity, we have cloned the cDNA and determined the primary structure of human and mouse ADAMTS-14, a novel ADAMTS displaying striking homologies with ADAMTS-2 and -3. The structure of the human gene, which maps to 10q21.3, and the mechanisms of generation of the various transcripts are described. The existence of two sites of initiation of transcription, in two different promoter contexts, suggests that transcripts resulting from these two sites can be differently regulated. The tissue distribution of ADAMTS-14, the regulation of the gene expression by various cytokines and the activity of the recombinant enzyme are evaluated. The potential function of ADAMTS-14 as a physiological aminoprocollagen peptidase *in vivo* is discussed.

PMID: 11741898 [PubMed - indexed for MEDLINE]

Cloning, expression analysis, and structural characterization of seven novel human ADAMTSs, a family of metalloproteinases with disintegrin and thrombospondin-1 domains.

Cal S, Obaya AJ, Llamazares M, Garabaya C, Quesada V, Lopez-Otin C.

Departamento de Bioquímica y Biología Molecular, Facultad de Medicina, Instituto Universitario de Oncología, Universidad de Oviedo, 33006, Oviedo, Spain.

ADAMTS (A Disintegrin And Metalloproteinase domain, with ThromboSpondin type-1 modules) is a recently described family of zinc-dependent proteases which play important roles in a variety of normal and pathological conditions, including arthritis and cancer. In this work, we report the identification and cloning of cDNAs encoding seven new human ADAMTSs. These novel enzymes have been called ADAMTS-13, -14, -15, -16, -17, -18, and -19. All of them show a domain organization similar to that of previously characterized family members, consisting of a signal sequence, a propeptide, a metalloproteinase domain, a disintegrin-like domain, a cysteine-rich region, and a variable number of TS-1 repeats. Expression analysis revealed that these ADAMTS genes are mainly expressed in fetal tissues, especially in lung (ADAMTS14, ADAMTS16, ADAMTS17, ADAMTS18, and ADAMTS19), kidney (ADAMTS14, ADAMTS15, and ADAMTS16), and liver (ADAMTS13, ADAMTS15 and ADAMTS18). Reverse transcriptase--polymerase chain reaction analysis also revealed the expression of some of these new ADAMTSs in different human adult tissues, such as prostate (ADAMTS13, ADAMTS17, and ADAMTS18), and brain (ADAMTS13, ADAMTS16, ADAMTS17, and ADAMTS18). High levels of ADAMTSs transcripts were also observed in some tumor biopsies and cell lines, including osteosarcomas (ADAMTS19), melanoma and colon carcinoma cells (ADAMTS13). Chromosomal location analysis indicated that the seven identified ADAMTS genes are dispersed in the human genome mapping to 9q34, 10q21, 11q25, 5p15, 15q24, 16q23, and 5q31, respectively. According to these results, together with a comparative analysis of ADAMTSs in other eukaryotic organisms, we conclude that these enzymes, with at least 18 distinct members encoded within the human genome, represent an example of a widely expanded protease family during metazoan evolution.

PMID: 11867212 [PubMed - indexed for MEDLINE]

Biochim Biophys Acta. 2001 Dec 30;1522(3):221-5.

[Related Articles](#), [Links](#)

ELSEVIER SCIENCE
FULL-TEXT ARTICLE

Characterization of ADAMTS14, a novel member of the ADAMTS metalloproteinase family.

Bolz H, Ramirez A, von Brederlow B, Kubisch C.

Institut fur Hmgentik, Universitats-Klinikum Eppendorf, Hamburg, Germany.
bolz@uke.uni-hamburg.de

ADAMTS (a disintegrin-like and metalloproteinase domain with thrombospondin type 1 modules) proteins constitute a family of zinc metalloproteinases which target and process extracellular matrix proteins. We cloned and characterized a novel human ADAMTS gene, ADAMTS14, which is located on human chromosome 10q2. ADAMTS14 exhibits the characteristic multidomain structure of ADAMTS proteins including four thrombospondin modules and shows highest similarity to ADAMTS3 and ADAMTS2. By RT-PCR analysis we demonstrated that ADAMTS14 is expressed in human retina and also at low levels in adult brain, lung and placenta.

PMID: 11779638 [PubMed - indexed for MEDLINE]



EXHIBIT "E"

MegaBlast

MEGABLAST 1.2.3-Paracel [2001-11-20]

Reference:

Zheng Zhang, Scott Schwartz, Lukas Wagner, and Webb Miller (2000),
"A greedy algorithm for aligning DNA sequences",
J Comput Biol 2000; 7(1-2):203-14.

Database: Homo_sapiens.latestgp.masked.fa
33,840 sequences; 200,810,911,373 total letters

Query= LEX221seqid19
(3759 letters)

Sequences producing significant alignments:

	Score (bits)	E Value
AL355344.20.1.149490	882	0.0
AC007484.2.50228.57677	854	0.0
AC069538.10.1.169772	654	0.0
AC007484.2.23991.29266	654	0.0
AC007484.2.57778.65289	387	e-104
AC007484.2.112555.124636	314	3e-82
AC010216.8.1.110753	139	1e-29

>AL355344.20.1.149490
Length = 149490

Score = 882 bits (444), Expect = 0.0
Identities = 444/444 (100%)
Strand = Plus / Plus

Query: 80 cagagctgcacccctctggaaagctcagtgactatggtgtgacagtgcgcgcacag 139
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct: 147159 cagagctgcacccctctggaaagctcagtgactatggtgtgacagtgcgcgcacag 147218

Query: 140 actttcgggacgcttccctccacgtgggtctggccaggcagcgcgcctctgcaggga 199
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct: 147219 actttcgggacgcttccctccacgtgggtctggccaggcagcgcgcctctgcaggga 147278

Query: 200 gcatggtagtgacacgccacccacactaccacgcacactccaggcaccc 259
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct: 147279 gcatggtagtgacacgccacccacactaccacgcacactccaggcaccc 147338

Query: 260 gcagccctctgcacccaggaggacccctgtggcctggcagggtggggccactccctct 319
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct: 147339 gcagccctctgcacccaggaggacccctgtggcctggcagggtggggccactccctct 147398

Query: 320 acttcaatgtcactgtttcgaaaaggaactgcacttgcgcctgcggccaaatcgagg 379
||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct: 147399 acttcaatgtcactgtttcgaaaaggaactgcacttgcgcctgcggccaaatcgagg 147458

Query: 380 tggtagtgccaggatcctcagtggagtggcaggaggatttcggagctgttccggcagc 439

|||
Sbjct: 147459 tggttagtgccaggatcctcagtggagtggcaggaggatttcggagctgttccggcagc 147518

Query: 440 cttacggcaggagtgtgtacactggagggtcacactggaaatgcctggggcagctgttg 499
|||

Sbjct: 147519 cttacggcaggagtgtgtacactggagggtcacactggaaatgcctggggcagctgttg 147578

Query: 500 ccatcagcaactgtgacggattgg 523
|||

Sbjct: 147579 ccatcagcaactgtgacggattgg 147602

Score = 159 bits (80), Expect = 1e-35

Identities = 81/82 (98%)

Strand = Plus / Plus

Query: 1 atggctccactccgcgcgtgctgtcctacctgctgccttgcactgtgcgcctgcrc 60
|||

Sbjct: 145409 atggctccactccgcgcgtgctgtcctacctgctgccttgcactgtgcgcctgcgcc 145468

Query: 61 gccgcgggcagccggaccccag 82
|||

Sbjct: 145469 gccgcgggcagccggaccccag 145490

>AC007484.2.50228.57677

Length = 7450

Score = 854 bits (430), Expect = 0.0

Identities = 441/444 (99%), Gaps = 3/444 (0%)

Strand = Plus / Minus

Query: 80 cagagctgcacctcttgaaaagctcagtgactatggtgtgacagtgcctgcagcacag 139
|||

Sbjct: 5458 cagagctgcacctcttgaaaagctcagtgactatggtgtgacagtgcctgcagcacag 5399

Query: 140 acttcggggacgcttcctctccacgtgggtctggccacgcgcgcctctgcaggga 199
|||

Sbjct: 5398 acttcggggacgcttcctctccacgtgggtctggccacgcgcgcctctgcaggga 5339

Query: 200 gcatggtagtgacacgcacccacactaccacgcacactccagtcacccgggtggctc 259
|||

Sbjct: 5338 gcatggtagtgacacgcacccacactaccacgcacactccagtcacccgggtggctc 5279

Query: 260 gcagccctctgcacccaggaggacccctgtggcctggcagggtggggccactccctct 319
|||

Sbjct: 5278 gcagccctctgcacccaggaggacccctgtggcctggcagggtggggccactccctct 5219

Query: 320 acttcaatgtcactgtttcggaaggaaactgcacttgccctgcggccaatcgaggt 379
Sbjct: 5218 acttcaatgtcactgtttcggaaggaaactgcacttgcg--tcggccaatcgaggt 5161

Query: 380 tggtagtgcaggatcctcagtggagtggcaggaggatttcggagctgtccggcagc 439
Sbjct: 5160 tggtagtgcaggatcctcagtggagtggcaggaggatttcggagctgtccggcagc- 5102

Query: 440 cttacggcaggagtgtgtacactggaggtgtcactggaatgcctgggcagctgttg 499
Sbjct: 5101 cttacggcaggagtgtgtacactggaggtgtcactggaatgcctgggcagctgttg 5042

Query: 500 ccatcagcaactgtgacggattgg 523
Sbjct: 5041 ccatcagcaactgtgacggattgg 5018

Score = 133 bits (67), Expect = 7e-28
Identities = 79/83 (95%), Gaps = 3/83 (3%)
Strand = Plus / Minus

Query: 1 atggctccactccgcgcgtgtgccttacctgtgccttcgc-actgtgcgcctgcrc 59
Sbjct: 7207 atggctccactccgcgcgtgtgccttacctgtgccttcgcactgt-cgctctgcgc 7149

Query: 60 cggccggcggcagccggaccccag 82
Sbjct: 7148 c-ccgcggcagccggaccccag 7127

>AC069538.10.1.169772
Length = 169772

Score = 654 bits (329), Expect = 0.0
Identities = 336/338 (99%), Gaps = 1/338 (0%)
Strand = Plus / Minus

Query: 3337 tcactgcccccccttcactcctggaaagcccattaccaggacccaggaccctgcagat 3396
Sbjct: 132825 tcactgcccccccttcactcctggaaagcccattaccaggacccaggaccctgcagat 132766

Query: 3397 gctgcagagcctcctggaaagccaacggatcagaggaccatcagcatggccgagccaca 3456
Sbjct: 132765 gctgcagagcctcctggaaagccaacggatcagaggaccatcagcatggccgagccaca 132706

Query: 3457 cagctcccaggagctctggataacaagctccccagggaccaggcatcccttgccttgc 3516
Sbjct: 132705 cagctcccaggagctctggataacaagctccccagggaccaggcatcccttgccttgc 132646

Query: 3517 acaccaatccctggagcatcctggagcatctccctaccaccccccgggggtgccttgg 3576
Sbjct: 132645 acaccaatccctggagcatcctggagcatctccctaccaccccccgggggtgccttgg 132586

Query: 3577 ggctggactcagacacacctacgccagtccctgaggacaaaggcaacctggagaagacctg 3636
Sbjct: 132585 ggctggactcagacacacctacgccagtccctgaggacaaaggcaacctggagaagacctg 132526

Query: 3637 agacatcccgaccagccctccctgctgacctgccccgg 3674
Sbjct: 132525 agacatcccgaccagccctccctgctg-cctccccgg 132489

Score = 405 bits (204), Expect = e-110
Identities = 207/208 (99%)
Strand = Plus / Minus

Query: 2730 gtgggtgacggaggagtgggtgcctgcagccggagctgtggaaagctgggggtgcagac 2789
Sbjct: 139531 gtgggtgacggaggagtgggtgcctgcagccggagctgtggaaagctgggggtgcagac 139472

Query: 2790 acggggatacagtgcctgctgcccttcataatggaaacccacaaggcatgccggccaa 2849
Sbjct: 139471 acggggatacagtgcctgctgcccttcataatggaaacccacaaggcatgccggccaa 139412

Query: 2850 agcctgcggggaccggcctgaggccgcacggccctgtctccgagtgcctgcggcagc 2909
Sbjct: 139411 agcctgtgcggggaccggcctgaggccgcacggccctgtctccgagtgcctgcggcagc 139352

Query: 2910 ccagtggaggctggagccctggcc 2937
Sbjct: 139351 ccagtggaggctggagccctggcc 139324

Score = 340 bits (171), Expect = 5e-90
Identities = 171/171 (100%)
Strand = Plus / Minus

Query: 2427 ggctctcccccaactgagggtggcccccgcagcagcctggctacaagtacgtcatcca 2486
Sbjct: 141854 ggctctcccccaactgagggtggcccccgcagcagcctggctacaagtacgtcatcca 141795

Query: 2487 tgaggacctgctgcccttatcggagcaacaatgtgctcctggaggatggacaccta 2546
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 141794 tgaggacctgctgcccttatcggagcaacaatgtgctcctggaggatggacaccta 141735

Query: 2547 tgagtggcgctcaagagctggccctgcagcaaggcctgtggaggagg 2597
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 141734 tgagtggcgctcaagagctggccctgcagcaaggcctgtggaggagg 141684

Score = 338 bits (170), Expect = 2e-89
Identities = 173/176 (98%)
Strand = Plus / Minus

Query: 1749 cccagcctatggaggccgcccyygtgcttaggcccattttcgagttaccaggctgcacag 1808
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 152344 cccagcctatggaggccgcccgtgcttaggcccattttcgagttaccaggctgcacag 152285

Query: 1809 cgaggagtgcctggacacctacgaggacttccggccagcagtgtgccaagcgcaactc 1868
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 152284 cgaggagtgcctggacacctacgaggacttccggccagcagtgtgccaagcgcaactc 152225

Query: 1869 stactatgtgcaccagaatgccaagcacagstgggtgccctacgagcctgacgatg 1924
||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 152224 ctactatgtgcaccagaatgccaagcacagctgggtgccctacgagcctgacgatg 152169

Score = 330 bits (166), Expect = 4e-87
Identities = 166/166 (100%)
Strand = Plus / Minus

Query: 2263 gtggtaagaaccaggtaaccggcagcttcattcctcaaccccaaggcaaggaagccaca 2322
||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 143510 gtggtaagaaccaggtaaccggcagcttcattcctcaaccccaaggcaaggaagccaca 143451

Query: 2323 agccggaccttcaccggcatggcctggagtgggaggatgcggtgaggatgccaaggaa 2382
||| ||| ||| ||| ||| ||| |||
Sbjct: 143450 agccggaccttcaccggcatggcctggagtgggaggatgcggtgaggatgccaaggaa 143391

Query: 2383 agcctcaagaccagcggccctgcctgaagccattgccatcctgg 2428
||| ||| ||| ||| ||| ||| |||
Sbjct: 143390 agcctcaagaccagcggccctgcctgaagccattgccatcctgg 143345

Score = 306 bits (154), Expect = 7e-80
Identities = 158/159 (99%), Gaps = 1/159 (0%)
Strand = Plus / Minus

Query: 952 cagtccctgagcctgatcgagcgcggaaaccctcacgcagcctggagcaggtgtgtcgc 1011
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 163232 cagtccctgagcctgatcgagcgcggaaaccctcacgcagcctggagcaggtgtgtcgc 163173

Query: 1012 tggcacactcccagcagcgcaggacccagccacgctgagcaccatgaccacgttgc 1071
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 163172 tggcacactcccagcagcgcaggacccagccacgctgagcaccatgaccacgttgc 163113

Query: 1072 ttccctacccggcaggactttggccctcagggtatgca 1110
||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 163112 ttccctacccggcaggactttggccctcagg-tatgca 163075

Score = 300 bits (151), Expect = 4e-78
Identities = 151/151 (100%)
Strand = Plus / Minus

Query: 1598 agtggtgcttcaaaggtaactgcatttggaaagtgcggagcagacatatggcaggatg 1657
||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 154491 agtggtgcttcaaaggtaactgcatttggaaagtgcggagcagacatatggcaggatg 154432

Query: 1658 gaggctggagctctggaccaagttgggtcatgttcgcgtcatgtggggcggtgc 1717
||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 154431 gaggctggagctctggaccaagttgggtcatgttcgcgtcatgtggggcggtgc 154372

Query: 1718 gatcccgagccggagctgcaacaaccctc 1748
||| ||| ||| ||| ||| |||
Sbjct: 154371 gatcccgagccggagctgcaacaaccctc 154341

Score = 286 bits (144), Expect = 6e-74
Identities = 144/144 (100%)
Strand = Plus / Minus

Query: 1209 gctcgcatggagcatgacggtcagggaaatggctgtgcagatgagaccgcctggcag 1268
||| ||| ||| ||| ||| ||| |||
Sbjct: 159446 gctcgcatggagcatgacggtcagggaaatggctgtgcagatgagaccgcctggcag 159387

Query: 1269 cgtagatggccccctggtcaggctgccttcaccgcattggccgcgtcagcaa 1328
||| ||| ||| ||| ||| |||
Sbjct: 159386 cgtagatggccccctggtcaggctgccttcaccgcattggccgcgtcagcaa 159327

Query: 1329 gctggagctcagccgtacccccc 1352
||| ||| ||| ||| |||
Sbjct: 159326 gctggagctcagccgtacccccc 159303

Score = 274 bits (138), Expect = 2e-70
Identities = 138/138 (100%)
Strand = Plus / Minus

Query: 2595 agggatccagttcaccaaatacggtgcggcgcagacgagaccacatggtcagcg 2654
|||||||
Sbjct: 141238 agggatccagttcaccaaatacggtgcggcgcagacgagaccacatggtcagcg 141179

Query: 2655 acacctgtgtgaccacaagaagaggccaaagcccattccgcggcgctgcaaccacacc 2714
|||||||
Sbjct: 141178 acacctgtgtgaccacaagaagaggccaaagcccattccgcggcgctgcaaccacacc 141119

Query: 2715 gtgccttcagcctgtgt 2732
|||||||
Sbjct: 141118 gtgccttcagcctgtgt 141101

Score = 264 bits (133), Expect = 2e-67
Identities = 134/135 (99%)
Strand = Plus / Minus

Query: 1924 gacgcccagaagtgtgagctgatctgccagtcggcggacacrgggacgtgggttcatg 1983
|||||||
Sbjct: 149784 gacgcccagaagtgtgagctgatctgccagtcggcggacacggggacgtgggttcatg 149725

Query: 1984 aaccaggtggttcacgatggacacgctgcagctaccggaccatacagcgtctgtcg 2043
|||||||
Sbjct: 149724 aaccaggtggttcacgatggacacgctgcagctaccggaccatacagcgtctgtcg 149665

Query: 2044 cgtggcgagtggtgt 2058
|||||||
Sbjct: 149664 cgtggcgagtggtgt 149650

Score = 264 bits (133), Expect = 2e-67
Identities = 133/133 (100%)
Strand = Plus / Minus

Query: 1353 ctcctacgactgcctcctcgatgaccctttgatcctgcctggccccagccccagagct 1412
|||||||
Sbjct: 158162 ctcctacgactgcctcctcgatgaccctttgatcctgcctggccccagccccagagct 158103

Query: 1413 gcctggatcaactactcaatggatgagcagtgccgcttgcattttggcagtggctacca 1472
|||||||
Sbjct: 158102 gcctggatcaactactcaatggatgagcagtgccgcttgcattttggcagtggctacca 158043

Query: 1473 gacctgcttggca 1485
|||||||
Sbjct: 158042 gacctgcttggca 158030

Score = 262 bits (132), Expect = 1e-66
Identities = 133/134 (99%)
Strand = Plus / Minus

Query: 2935 cagtgcctgtccacctgtggagagggcatccaggcagcggcaggtggtgtgcaggaccAAC 2994
|||||||
Sbjct: 135372 cagtgcctgtccacctgtggagagggcatccaggcagcggcaggtggtgtgcaggaccAAC 135313

Query: 2995 gccAACAGCCTCGGGCATTGCGAGGGGGATAGGCCAGACACTGTCCAGGTCTGCARCCCTG 3054
|||||||
Sbjct: 135312 gccAACAGCCTCGGGCATTGCGAGGGGGATAGGCCAGACACTGTCCAGGTCTGCAGCCTG 135253

Query: 3055 CCCGCCTGTGGAGG 3068
|||||||
Sbjct: 135252 CCCGCCTGTGGAGG 135239

Score = 254 bits (128), Expect = 2e-64
Identities = 128/128 (100%)
Strand = Plus / Minus

Query: 3179 tgtgtgcagcgagccctgcacggagacaggctgtttctgccAGATGGAAGTGTCTG 3238
|||||||
Sbjct: 132983 tgtgtgcagcgagccctgcacggagacaggctgtttctgccAGATGGAAGTGTCTG 132924

Query: 3239 atcgctactgtccattcccgctaccaccggctctgtgtgtcctgcATCAAGAAGG 3298
|||||||
Sbjct: 132923 atcgctactgtccattcccgctaccaccggctctgtgtgtcctgcATCAAGAAGG 132864

Query: 3299 CCTCGGGC 3306
|||||||
Sbjct: 132863 CCTCGGGC 132856

Score = 250 bits (126), Expect = 4e-63
Identities = 126/126 (100%)
Strand = Plus / Minus

Query: 2058 gcctgtcggctgtgacaaggaggtgggtccatgaaggcggatgacaagtgtggagtctg 2117
|||||||

Sbjct: 149266 gcctgtcggctgtgacaaggaggtgggtccatgaaggcgatgacaagtgtggagtctg 149207

Query: 2118 cgggggtgacaactcccactgcaggactgtgaaggggacgctggcaaggcctccaagca 2177
|||||||

Sbjct: 149206 cgggggtgacaactcccactgcaggactgtgaaggggacgctggcaaggcctccaagca 149147
|||||||

Query: 2178 ggcagg 2183
||||||

Sbjct: 149146 ggcagg 149141

Score = 221 bits (111), Expect = 4e-54

Identities = 112/113 (99%)

Strand = Plus / Minus

Query: 3066 agaaaatcaccagaactccacggtgagggccatgtctggaaacttggacgccagaggg 3125
|||||||

Sbjct: 135158 agaaaatcaccagaactccacggtgagggccatgtctggaaacttggacgccagaggg 135099
|||||||

Query: 3126 gcagtgggtgccacaatctgraccctacatcccattaacaagatatcatcaa 3178
|||||||

Sbjct: 135098 gcagtgggtgccacaatctgaaccctacatcccattaacaagatatcatcaa 135046
|||||||

Score = 219 bits (110), Expect = 1e-53

Identities = 112/114 (98%)

Strand = Plus / Minus

Query: 1486 ttcaaggacctttgagccctgcaaggcagctgtggtcagccatcctgacaaccmgtayttc 1545
|||||||

Sbjct: 156651 ttcaaggacctttgagccctgcaaggcagctgtggtcagccatcctgacaaccgtacttc 156592
|||||||

Query: 1546 tgcaagaccaagaaggggccccctggatggactgagtgtgcacccggcaag 1599
|||||||

Sbjct: 156591 tgcaagaccaagaaggggccccctggatggactgagtgtgcacccggcaag 156538
|||||||

Score = 219 bits (110), Expect = 1e-53

Identities = 110/110 (100%)

Strand = Plus / Minus

Query: 1100 cagggtatgcacccgtcactggcatgtgtcacccctgaggagctgtgcacccatg 1159
|||||||

Sbjct: 161080 cagggtatgcacccgtcactggcatgtgtcacccctgaggagctgtgcacccatg 161021
|||||||

Query: 1160 aggatggcttcctcagccttcgtatagctcatgagacccggcacgtg 1209

Sbjct: 161020 |||||||aggatggcttctcctcagccttcgtatgcatacgaccggcacgtg 160971

Score = 167 bits (84), Expect = 5e-38
Identities = 84/84 (100%)
Strand = Plus / Minus

Query: 871 gtagatgagatttaccacgatgagtccctgggggtcatataaatattgccctcgccgc 930
Sbjct: 164037 |||||||gtagatgagatttaccacgatgagtccctgggggtcatataaatattgccctcgccgc 163978

Query: 931 ttgatcatgggtggctaccgacag 954
Sbjct: 163977 |||||||ttgatcatgggtggctaccgacag 163954

Score = 161 bits (81), Expect = 3e-36
Identities = 81/81 (100%)
Strand = Plus / Minus

Query: 2178 ggcaggagctctcaagctggtgcaagatcccagcaggtgccaggcacatccagattgaggc 2237
Sbjct: 148110 |||||||ggcaggagctctcaagctggtgcaagatcccagcaggtgccaggcacatccagattgaggc 148051

Query: 2238 actggagaagtccccccacccg 2258
Sbjct: 148050 |||||||actggagaagtccccccacccg 148030

Score = 145 bits (73), Expect = 2e-31
Identities = 73/73 (100%)
Strand = Plus / Minus

Query: 3687 gccctgccatcccactggcacgtttacactctgtgtactgccccgtgactccagctcg 3746
Sbjct: 132475 |||||||gccctgccatcccactggcacgtttacactctgtgtactgccccgtgactccagctcg 132416

Query: 3747 aggacacacatag 3759
Sbjct: 132415 |||||||aggacacacatag 132403

Score = 89.9 bits (45), Expect = 1e-14
Identities = 66/73 (90%)
Strand = Plus / Minus

Query: 1375 gaccctttgatcctgcctggcccccagccccagagctgcctggatcaactactcaatg 1434
Sbjct: 158909 gaccctttgagcccacctggcccccagccccagagctgcctggatcgacttctcaatg 158850

Query: 1435 gatgagcagtgcc 1447
Sbjct: 158849 gatgaacagtgcc 158837

>AC007484.2.23991.29266
Length = 5276

Score = 654 bits (329), Expect = 0.0
Identities = 336/338 (99%), Gaps = 1/338 (0%)
Strand = Plus / Minus

Query: 3337 tcactgcccccttctccactcctggaaagcccccattaccaggacccaggaccctgcagat 3396
Sbjct: 2310 tcactgcccccttctccactcctggaaagcccccattaccaggacccaggaccctgcagat 2251

Query: 3397 gctgcagagcctcctggaaagccaacggatcagaggaccatcagcatggcgagccaca 3456
Sbjct: 2250 gctgcagagcctcctggaaagccaacggatcagaggaccatcagcatggcgagccaca 2191

Query: 3457 cagctcccaggagctctggataacaagctccccaggaccatcccttgccttgag 3516
Sbjct: 2190 cagctcccaggagctctggataacaagctccccaggaccatcccttgccttgag 2131

Query: 3517 acaccaatccctggagcatctggagcatctccctaccaccccccgggggctgccttgg 3576
Sbjct: 2130 acaccaatccctggagcatctggagcatctccctaccaccccccgggggctgccttgg 2071

Query: 3577 ggctggactcagacacacctacgcccagtccctgaggacaaaggcaacctggagaagacctg 3636
Sbjct: 2070 ggctggactcagacacacctacgcccagtccctgaggacaaaggcaacctggagaagacctg 2011

Query: 3637 agacatcccgaccaggcccttcgtgacctggccgg 3674
Sbjct: 2010 agacatcccgaccaggcccttcgtg-cctcccccgg 1974

Score = 262 bits (132), Expect = 1e-66
Identities = 133/134 (99%)
Strand = Plus / Minus

Query: 2935 cagtgcctgtccaccgtggagagggcatccagcagcggcaggtggtgtgcaggaccaac 2994

Sbjct: 4860 ||||| cagtgc||tcgc||acc||tggagagggc||atcc||cagc||cg||cg||cagg||tg||gtgc||aggac||aa||ac 4801

Query: 2995 gccaacagcctcgccattgcggataggccagacactgtccaggctgcacccctg 3054
Sbjct: 4800 gccaacagcctcgccattgcggataggccagacactgtccaggctgcacccctg 4741

Query: 3055 cccgcctgtggagg 3068
| | | | | | | | | | | |
Sbjct: 4740 cccgcctgtggagg 4727

Score = 254 bits (128), Expect = 2e-64
Identities = 128/128 (100%)
Strand = Plus / Minus

Query: 3179 tgtgtgcagcggagccctgcacggagacaggctgtcttctgccagatggaagtgtcg 3238
|||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Sbjct: 2468 tgtgtgcagcggagccctgcacggagacaggctgtcttctgccagatggaagtgtcg 2409

Query: 3239 atcgctactgctccattcccggtaccacccggctctgctgtgtgtcctgcatcaagaagg 3298
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 2408 atcgctactgctccattcccggtaccacccggctctgctgtgtgtcctgcatcaagaagg 2349

Query: 3299 cctcgccg 3306
|||
Sbjct: 2348 cctcgccg 2341

Score = 221 bits (111), Expect = 4e-54
Identities = 112/113 (99%)
Strand = Plus / Minus

Query: 3066 agggaaatcaccagaactccacggtgagggccgatgtctggaaacttgggacgccagaggg 3125
Sbjct: 4646 agggaaatcaccagaactccacggtgagggccgatgtctggaaacttgggacgccagaggg 4587

Query: 3126 gcagtgggtgccacaatctgracccctacatcccattaaacaagatatacatcaa 3178
Sbjct: 4586 gcagtgggtgccacaatctgaacccctacatcccattaaacaagatatacatcaa 4534

Score = 145 bits (73), Expect = 2e-31
Identities = 73/73 (100%)
Strand = Plus / Minus

Query: 3687 gccctgccatcccactggcacgttacactctgtgtactgccccgtgactccagctcag 3746
Sbjct: 1960 gccctgccatcccactggcacgttacactctgtgtactgccccgtgactccagctcag 1901

Query: 3747 aggacacacata 3759
Sbjct: 1900 aggacacacata 1888

>AC007484.2.57778.65289
Length = 7512

Score = 387 bits (195), Expect = e-104
Identities = 195/195 (100%)
Strand = Plus / Minus

Query: 678 agccttggcctggagacacctccaaacctgctggcctgggtggggaccagctggcga 737
Sbjct: 2978 agccttggcctggagacacctccaaacctgctggcctgggtggggaccagctggcga 2919

Query: 738 cacagagcggaaagcggcgcatgccaagccaggcagctacagcatcgagggtgcgtgg 797
Sbjct: 2918 cacagagcggaaagcggcgcatgccaagccaggcagctacagcatcgagggtgcgtgg 2859

Query: 798 ggtggacgactcggtggttcgcttcatggcaaggagcatgtgcagaactatgtcctcac 857
Sbjct: 2858 ggtggacgactcggtggttcgcttcatggcaaggagcatgtgcagaactatgtcctcac 2799

Query: 858 cctcatgaatatcgt 872
Sbjct: 2798 cctcatgaatatcgt 2784

>AC007484.2.112555.124636
Length = 12082

Score = 314 bits (158), Expect = 3e-82
Identities = 158/158 (100%)
Strand = Plus / Plus

Query: 522 ggcgggcctcatccgcacagacagcaccgacttcttcattgagcctctggagcggggcca 581
Sbjct: 10474 ggcgggcctcatccgcacagacagcaccgacttcttcattgagcctctggagcggggcca 10533

Query: 582 gcaggagaaggaggccagcggaggacacatgtggtgtaccgcggaggccgtccagca 641
Sbjct: 10534 gcaggagaaggaggccagcggaggacacatgtggtgtaccgcggaggccgtccagca 10593

Query: 642 ggagtgggcagaacctgacgggacctgcacaatgaag 679
Sbjct: 10594 ggagtgggcagaacctgacgggacctgcacaatgaag 10631

>AC010216.8.1.110753
Length = 110753

Score = 139 bits (70), Expect = 1e-29
Identities = 101/112 (90%), Gaps = 4/112 (3%)
Strand = Plus / Plus

Query: 1490 ggacctttgagccctgcaaggcagctgtggtcagccatcctgacaaccmgtayttctgca 1549
Sbjct: 32976 ggacctttgacccctgcaaggcagctgtggtcagccatcctgacaaccctactttgca 33035

Query: 1550 agaccaagaaggggcccccgct-ggatggactga-gtgtgcacccggcaag 1599
Sbjct: 33036 agaccaagaaggggcccc-cttgacggact-atgtgtgcacctggcaag 33085

Score = 129 bits (65), Expect = 1e-26
Identities = 83/89 (93%)
Strand = Plus / Plus

Query: 1261 ctgggcagcgtcatggcgccctggtcaggctgccttccaccgcttccattggtcccgc 1320
Sbjct: 31087 ctgggcagcatcatggcgccctggtcaggccgccttccaccgcttccactggtcccgc 31146

Query: 1321 tgcaagctggagctcagccgctacct 1349
Sbjct: 31147 tgcaagccagcaggagctgagccgctacct 31175